

Rationology 101

How the Author of Genesis Got It Right (and the Golden Rule Got It Wrong)

Phil Roberts, Jr.
philrob@ix.netcom.com

Abstract: It is generally taken as a given that rationality is strictly a matter of adjudicating means to ends. Based on the premise that 'feelings of worthlessness' are a maladaptive byproduct of the evolution of rationality, I forego this convention by proposing a theory of rationality that encompasses the rationality of ends. One of the more interesting implications of this approach is that the moral maxim, 'Love (intrinsically value) your neighbor as you love (intrinsically value) yourself' can be construed as an imperative of an implicit theory of rationality in which 'being rational' is simply a matter of 'being objective'. Furthermore, by demonstrating how this implicit theory can address various rationality paradoxes (rational irrationality, epistemic vs. practical rationality conflict, the "rationality debate" [Stein, 1996], the Prisoners' Dilemma, etc.), its epistemic credentials can be shown to surpass those of competing theories such as the means/end theory, rational choice theory, egoism, utilitarianism, etc. In the final section of the paper I employ some of these insights to derive a moral 'ought' from an epistemic 'is'.

Background

In Roberts, 2005a, I have argued that the insistence that psychology be anchored in, if not necessarily restricted to, "the third-person point of view... [of] ...objective physical science" (Dennett, 1991, pp. 71-73) has led to serious errors of omission in the behavioral sciences.⁽¹⁾ In Roberts, 2005b, I attempted to address this imbalance by presenting an outline of a theory of ego/self-worth related emotion that began as follows:

Objective: To account for ego/self-worth related emotional need (e.g., needs for love, purpose, meaning, acceptance, attention, moral integrity, recognition, achievement, wealth, power, dignity, romance, modesty, fame, immortality, religion, autonomy, justice, etc.) and disorder (e.g., anxiety, depression, addiction, suicide, etc.) within the context of an evolutionary scenario; i.e., to synthesize natural science and the humanities; i.e., to answer the question: 'Why are the members of one particular species of naturally selected organism (*Homo Sapiens*) expending significant amounts of effort and energy on the biologically bizarre non-physical objective of maximizing self-worth?'

General Observation: The species in which rationality is most developed is also the one in which individuals have the greatest difficulty in maintaining an "adequate" sense of self-worth, often going to extraordinary lengths in doing so (e.g., Evel Knievel, celibate monks, 9/11 terrorists, etc.).

General Hypothesis: Rationality⁽²⁾ is antagonistic to psychocentric stability (i.e., maintaining an "adequate" sense of self-worth).

Explanation #1 (psychodynamics): In much the manner our rationality allows for the subordination of lower emotional concerns and values (pain, fear, anger, sex, etc.) to more global concerns (concern for the self as a whole), so too, these more global

concerns and values can themselves become reevaluated and subordinated to other more global, more objective considerations. And, if this is so, and assuming that emotional disorder emanates from a deficiency in self-worth resulting from precisely this sort of experientially based reevaluation, then it can reasonably be construed as a natural malfunction resulting from one's rational faculties functioning a tad too well.

Explanation #2 (rationality theory): Being the blind arational process that she is, Mother Nature instills in all her creatures a sense of their own importance (or of the importance of their needs) that is rationally inordinate. And, as a species reaches a certain stage in its rational/cultural/memetic development, its members increasingly come to question this inordinacy, and increasingly come to require reasons (justification) for maintaining it (needs for love, purpose, meaning, acceptance, etc.).

The implications of this conjecture are extensive, including implications with respect to emotional disorder (that it is a valuative affliction), indeterminism (that we are less valuatively/conatively determined than other species), incompleteness (that 'feelings of worthlessness' empirically vindicate the Lucas/Penrose perspective on Gödel's theorem) and ethics:

Ethics: Since, according to this explanation, more rational equates or correlates with more valuatively objective, the moral maxim, 'Love (i.e., intrinsically value) your neighbor as you love (intrinsically value) yourself' could be construed as an imperative of an implicit theory of rationality in which 'being rational' is simply a matter of 'being objective'. This would also mean that, to the extent this "implicit theory" turns out to be "true", the author of Genesis actually got it right in referring to our awareness of right and wrong as a form of knowledge.

This paragraph explains my reference to the author of Genesis in the subtitle. However, there is a problem with this explanation that I addressed in my included commentary on the outline:

Ethics: "the author of Genesis got it right": Although the Genesis reference to a knowledge of good and evil is compatible with my thesis that morality issues from an implicit theory of rationality, Adam and Eve's awareness of right and wrong was manifested, not in an awareness they should 'love others as they loved themselves' (valuative objectivity), but rather in an awareness of and discomfort with their nakedness which, on the surface, might seem incompatible with my thesis:

As they ate [the apple], suddenly they became aware of their nakedness, and were embarrassed. So they strung fig leaves together to cover themselves around the hips (Genesis 3:6).

This feature of our morality can be accommodated to my premise by simply assuming that our sexual appetites are as much an impediment to our valuative objectivity at the intrapersonal (prudential) level as selfishness is an impediment to our valuative objectivity at the interpersonal level. And, although this is a characteristic of most of our biological emotions (fear, anger, etc.) and appetites (hunger, thirst, etc.), sex is in a category all by itself. It's incoherent with the remainder of our biological functions in that it operates entirely in the interests of a few cells in the gonads rather than the interests of our physical well-being. This incoherence is further amplified by the fact that it requires an appreciation for regions and features of the body that in nonsexual circumstances we find the least aesthetically appealing. As such, like open displays of

pain, fear and anger in certain social contexts, explicit sexual values/impulses are a threat to our sense of self-control and our need to think of ourselves as rational, i.e., valuatively objective, beings.

If I am right about this, and our moral norms do indeed issue from a shared implicit theory in which ‘being rational’ is a simply matter of ‘being objective’, then we have reason to believe that the ‘Golden Rule’ may leave a bit to be desired as a moral maxim, a matter I have also commented on in my 2005 paper and which rounds out my explanation of the rationale for the subtitle of this paper:

Ethics: “Love your neighbor as you love yourself”: It is important not to confuse this maxim with the Golden Rule, i.e., “Do unto others as you would have them do unto you”. The Golden Rule is unsuitable as a moral maxim in that it can too easily be misconstrued as a prudential maxim in which one is advised to do nice things in the expectation that others will return the favor (reciprocity). Furthermore, it lends itself to the view that morality is more a matter of doing something than a matter of being something, i.e., valuatively objective and, as such, fails to do justice to the fact that morality is more a matter of selfless (i.e., valuatively objective) intention than a matter of overt behavior.

The Implicit Theory

1. The moral maxim, ‘Love (intrinsically value) your neighbor as you love (intrinsically value) yourself’ is of the form, ‘Given that one values X such and such an amount, then one ought to value Y such and such an amount’. As such, it is analogous to a cognitive syllogism of the form, ‘Given that X is true, then one ought to believe Y is true’, e.g., that Socrates is mortal, etc. In other words, the moral maxim makes no more claim about what is of value in the world than a cognitive syllogism lays claim to what is true in the world. It is strictly limited to the relationship between a valuative premise and its valuative conclusion in much the manner a cognitive syllogism is limited to the relationship between a cognitive premise and its cognitive conclusion. For the aesthetically minded, this apparent symmetry between the logic of the cognitive and valuative domains implicit in the moral maxim could itself be construed as an indication that our “shared implicit theory” (assuming that is what it is) exhibits a certain degree of elegance.
2. In addition to its apparent logical symmetry, our “shared implicit theory” also appears to have much in common with a longstanding paradigm for the rationality of ends, the formalization of prudence in the guise of the ‘equal weight’ criterion:

My feelings a year hence should be just as important to me as my feelings next minute, if only I could make an equally sure forecast of them. Indeed, this equal and impartial concern for all parts of one’s conscious life is perhaps the most prominent element in the common notion of the rational (Sidgwick, 1907).

Converting this paradigm to one that is compatible with the notion that ‘being rational’ is simply a matter of ‘being objective’ can be accomplished by simply abandoning the unwarranted assumption that ‘being rational’ necessarily entails ‘being self-interested’ on the grounds that there is no reason to assume that what is rational necessarily equates with what is adaptive:

Special concern for one's own future would be selected by evolution: Animals without such concern would be more likely to die before passing on their genes. Such concern would remain, as a natural fact, even if we decided that it was not justified. By thinking hard about the arguments, we might be able briefly to stun this natural concern. But it would soon revive... The fact that we have this attitude cannot therefore be a reason for thinking it justified. Whether it is justified [i.e., rational] is an open question, waiting to be answered (Parfit, 1984).

What you are left with is a theory in which the rationality of an end is in some manner a function of its objectivity and in which there is no fixed perimeter for where that objectivity should end.

3. Assuming this can reasonably be interpreted as a function of the comprehensiveness of the underlying considerations assessed from an objective perspective, one might then infer that a moral objective would qualify as relatively more rational than a prudential objective, a prudential objective would qualify as relatively more rational than imprudently responding to the whims of one's biological impulses (e.g., fear, anger, sex, etc.) with no regard for their overall consequences, and in which no concrete objective qualifies as rational in any but a relative sense of the term (i.e., no fixed perimeter).
4. This would also mean that our common sense ascriptions in which we refer to this or that as rational or irrational would have to be reinterpreted as shorthand for 'X compares favorably or unfavorably to the norm'.

Virtues of the Implicit Theory

1. The implicit theory can resolve the paradox of rational irrationality.

On pages 13 and 14 of *Reasons and Persons*, Derek Parfit offers a hypothetical scenario in which there is a significant likelihood that a robber will inflict grave harm on someone's family irrespective of whether the individual conforms to the robber's demands or not and in which, given the specific circumstances, far and away the best alternative would be to take a "special drug" that causes one to become temporarily irrational:

While I am in this state, I shall act in ways that are very irrational. There is a risk that, before the police arrive I may harm myself or my children. But, since I have no gun, this risk is small. And making myself irrational is the best way to reduce the great risk that this man will kill us all.

On any theory of rationality, it would be rational for me, in this case, to cause myself to become irrational. An acceptable theory about rationality can tell us to cause ourselves to do, what in its own terms, is irrational.

I would argue that what Parfit is actually consulting here is not any of the current theories of rationality, many of which would indeed sanction rational irrationality in the above scenario and thereby qualify as self-defeating (i.e., false), but rather a shared implicit theory in which 'being rational' is simply a matter of 'being objective', and in which no objective is rational in any but a relative sense of the term. This would explain how he could get away with asserting

a logical contradiction that none of us finds cognitively dissonant, in that underlying the absolutist terminology (rational vs. irrational) would be the shared understanding that in the above scenario the individual would actually be opting to become relatively less rational for a time as a means to a relatively more rational end (protecting his family), and in which the temporary reduction in rationality is merely an extension of the “irrationality” (lack of objectivity) that is part and parcel of fixating on a supremely valued end irrespective of the context.

2. The implicit theory can ameliorate epistemic vs. practical rationality conflict.

Rationality is widely believed to come in two basic flavors, an epistemic variety and a practical variety. The tension that can result when these two different types of rationality appear to collide has been engagingly explored in John Barth’s novels, and Steve Nathanson’s, *The Ideal of Rationality*. For example, a conflict between the epistemic and practical requirements arises on those occasions in which it is in one’s best interest to believe falsehood. Religious beliefs could quite possibly qualify in this regard, in which we might assume they are embraced because they bestow huge emotional benefits and increase group cohesion, albeit at the expense of having to abandon “the ethics of belief”:

It is wrong always, everywhere, for everyone, to believe anything upon insufficient evidence (W. K. Clifford).

Although perhaps not immediately evident, both the epistemic and the practical notions actually presuppose that rationality is a strategic attribute -- simplistically, as a matter of maximizing means to ends (e.g., the maximizing of true belief, etc.). However, in a theory of rationality in which ‘being rational’ is simply a matter of ‘being objective’, and in which there is no fixed perimeter for where that objectivity should end, there is no such thing as a rational end, only ends that are relatively more or less rational than others based on the comprehensiveness of their underlying considerations assessed from an objective perspective. As such, the tension that appears to exist between the two domains is not a tension in the nature of rationality itself, but rather a tension between two relatively rational objectives (epistemic vs. practical) in which an obsession with either to the exclusion of the other would be considered less rational than a more balanced and objective approach.

3. The implicit theory can resolve the “rationality debate” in a straightforward manner.

If our common sense ascriptions in which we refer to X as rational or irrational are indeed a shorthand for ‘X compares favorably or unfavorably to the norm’, as entailed by the implicit theory, then experimental evidence that we routinely violate established rules of deductive and probabilistic logic (e.g., Kahneman and Tversky, 1973) could not be construed as evidence that untutored adult humans are irrational, as scores of experts on the subject have maintained (‘Open Peer Commentary’ in Cohen, 1981). Nor would it be necessary to introduce a competence/performance distinction or to contemplate the implications of a wide versus a narrow reflective equilibrium in order to establish why this must be so (Cohen’s target article and ‘Author’s Response’; Stein, 1996; etc.). This is because, when translated into the sort of relativistic assertion presumed to underlie the absolutist language, an affirmative answer to questions such as ‘Can Human Irrationality Be Experimentally Demonstrated?’ (the title of the Cohen symposium) or ‘Could Man Be An Irrational Animal?’ (Stich, 1985) would amount to asserting a logical contradiction (e.g., ‘Can Subnormal Rationality Be Experimentally Demonstrated To Be the Norm?’, etc.).

4. The implicit theory is immune to the paradox of ‘The Prisoners’ Dilemma’.

A veritable mountain of literature has been written on a well known paradox in game theory (e.g., Axelrod, 1984) in which two prisoners will receive a light sentence if neither one rats on the other, both will receive lengthy sentences if both rat, but in which no matter what the other prisoner does, each prisoner will be individually better off by ratting. In this variation on the theme of “the tragedy of the commons”, the dictates of game theoretic rationality prescribe a course of action for each player that condemns both to lengthy sentences which, for many theoreticians, has seemed counterintuitive, and perhaps even paradoxical. In contrast, in the world of the “implicit theory” rationality prescribes, not action, but values, and in which the more rational the players become the more likely they will exhibit genuine concern for each other and, presumably, cooperate to their mutual benefit. Apparently, in the world of the implicit theory, there would be nothing but peace and harmony and happy endings.

5. The implicit theory is compatible with our common sense understanding of jury selection.

Unlike many current theories of rationality (e.g., means/end theory, rational choice theory, etc.), the implicit theory does not entail the counterintuitive conclusion that relatives of the accused should be excluded from jury duty because they would be too rational (e.g., too efficient or strategically logical in achieving their desired end of helping a relative), but rather because they would be less rational (less valuatively objective) than non-related jurors.

Derivation of a Moral ‘Ought’

There is a notorious logical gap between ‘is’ statements and ‘ought’ statements (e.g., Hume, 1739), with those who attempt to derive the latter from the former often criticized for committing what has come to be known as ‘the naturalistic fallacy’ (e.g., Tullberg and Tullberg, 2001). However, if our moral ‘oughts’ do indeed issue from an implicit theory of rationality in which ‘being rational’ is simply a matter of ‘being objective’, then those ‘oughts’ can actually be epistemically justified in the following manner:

1. Assume that ‘being rational’ is not simply a matter of...

being efficient (means/end theory),
being logical (computationalism),
being self-interested (egoism),
fulfilling one’s desires (hedonism),
maximizing true belief (e.g., Clifford),
being practical (pragmatism),
being strategically logical (rational choice theory, etc.),
conformity to a universalizable maxim (Kant),
maximizing global well-being (utilitarianism),
truth and falsehood (Hume),
etc.,

but simply a matter of...

‘being objective’, not only cognitively, but valuatively as well (impartiality), and in which no individual, belief, theory, objective, etc. is rational/objective in any but a relative sense of the term.

2. Corroborate the epistemic credentials of the above theory/definition by demonstrating that it can maximize explanatory coherence better than any of its competitors (e.g., the means/end theory, rational choice theory, etc.).

For example, the above theory/definition...

- can explain what morality “is” (p. 2),
- can unify the sexual and altruistic aspects of morality (pp. 2,3),
- exhibits logical symmetry between the valuative and cognitive domains (p. 3),
- has much in common with a longstanding paradigm for rationality (pp. 3,4),
- can resolve the paradox of rational irrationality (pp. 4,5),
- can ameliorate cognitive vs. practical rationality conflict (p. 5),
- can resolve the “rationality debate” in a straightforward manner (p. 5),
- is immune to the paradox of the Prisoner’s Dilemma (p. 6),
- is compatible with our common sense understanding of jury selection (p. 6),
- can afford an avenue for the justification of a moral ‘ought’ (pp. 6,7),
- etc.

3. Derive the ‘ought’ component of the is/ought derivation via the syllogism...

Premise:

Given that one is rational,

Connective:

then one ‘ought’ to...

Moral Maxim:

‘Love (intrinsically value) one’s neighbor as one loves (intrinsically values) oneself’, i.e., one ‘ought’ to be valuatively objective (impartial).

Conclusion

The premise that ‘being rational’ is simply a matter of ‘being objective’ is an implication of the theory of emotion presented in Roberts, 2005b. I will leave it to others to decide if the explanatory coherence itemized in step two above is sufficient to warrant that the source of that coherence be allowed to join the ranks of those theories we fallible human beings refer to as “true”.

Footnotes

1. Gleitman’s *Psychology* (1981), at one time the bible according to cognitive science, offers over 700 pages on everything from acquisition curves to zygotes, without a single reference to self-esteem, feelings of worthlessness, guilt, self-worth, etc. And, Eibl-Eibesfeldt’s sizeable tome on *Human Ethology* (1989) somehow manages 850 pages on topics ranging from abstraction to zebra finches, while studiously ignoring this same class of features. The same applies to *The Adapted Mind* (Barkow, Cosmides and Tooby, 1992), viewed by many as the cornerstone of evolutionary psychology.

One of the characteristics of the majority of modern psychological theories, aside from the arbitrariness of so many of their claims, is their frequently ponderous irrelevance. The cause, both of the irrelevance and of the arbitrariness, is the evident belief of their exponents that one can have a science of human nature while consistently ignoring man’s most significant and distinctive attributes (Nathaniel Branden).

2.

The concept of rationality, one might say, is incorrigibly elusive... I believe it is fair to say that in philosophical discussions of rationality, there is a sense in which we do not know what we are talking about and can never do so, if what is demanded is a concise definition (Max Black).

If one assumes, as I have conjectured, that feelings of worthlessness are a byproduct of the evolution of rationality in the sense of resulting from a more comprehensive and objective understanding of how the world is put together (Roberts, 2005b), it seems to follow that what Stein (1996) has referred to as the “standard picture of rationality” must be mistaken -- that ‘being rational’ is not so much a matter of slavishly conforming to established rules of inference (a process) as a matter of ‘being able to “see” what is going on’ as a result of reasoning that has already transpired, whether one’s own or culturally acquired (the product of a process). Accordingly, I assume our common sense rationality assessments can reasonably be construed as appraisals of a mental map of sorts in which the cognitive component of this “seeing” correlates with the extent to which the map is comprised of beliefs that accurately and coherently represent reality including, among many other things, beliefs about how to acquire beliefs that accurately and coherently represent reality (reflected in how well one reasons).

Compatible with the foregoing and the implication that rationality is a matter of degree, this would also mean that when we refer to an individual as "rational" or "irrational" that we are simply expressing a rough appraisal of how this individual's mental map compares to the norm and that the failure to appreciate this quirk in our ordinary use of words has probably been at the core of a considerable amount of confusion and disagreement (e.g., Cohen, 1981). For this reason, I do not construe experimental evidence that humans routinely violate established rules of inference (e.g., Kahneman and Tversky, 1973) as evidence that humans are irrational, as some have seemed to suggest (e.g., Tversky and Kahneman, 1974; Nisbett and Borgida, 1975; Slovic et al. 1976; etc.), but rather as evidence that expert opinion might be relatively more rational than the norm (in terms of ‘being able to “see” what is going on’) where such matters are concerned in much the manner some might argue that expert opinion has been relatively less rational than the norm with regard to its longstanding love affair with “the standard picture.”*

One of the slipperiest terms in the philosophical lexicon, 'rationality' is many things to many people (Alvin Plantinga).

*As is so often the case, I suspect the lover in question may have been blinded by lust, on this occasion, the lust to reduce mind to matter via the reduction of rationality to logic/ rules/ processes/ principles/ procedures etc. that can be instantiated in a computer.

References

1. Axelrod, Robert (1984), *The Evolution of Cooperation*.
2. Barkow, Cosmides and Tooby ed. (1992), *The Adapted Mind*.
3. Barth, John (1957), *The Meaning of Life*.
4. Barth, John (1972), *The Floating Opera*.
5. Barth, John (1969), *The End of The Road*.
6. Carol, John and Payne, John (1976), *Cognition and Social Behavior*.

7. Clifford, W. K. (1879), 'The Ethics of Belief', *Lectures and Essays*, vol. 2.
8. Cohen, L. Jonathan (1981), 'Can Human Irrationality Be Experimentally Demonstrated', *Behavioral and Brain Sciences*, 4, 317-370; 6, 487-533; 10, 311-13.
9. Dennett, D. C. (1991), *Consciousness Explained*.
10. Eibl-Eibesfeldt, Irenaus (1989), *Human Ethology*.
11. Gleitman, Henry (1981), *Psychology*.
12. Hume, David (1739), *A Treatise of Human Nature*.
13. Kahneman, D. and Tversky, A. (1973), 'On the Psychology of Prediction', *Psychological Review*, 80, 237-51.
14. Nathanson, Stephen (1985), *The Ideal of Rationality*.
15. Nisbett, Richard and Borgida, Eugene (1976), 'Attribution and the Psychology of Prediction', *Journal of Personal and Social Psychology*, 32, p. 935).
16. Parfit, Derek (1984), *Reasons and Persons*.
17. Roberts, Jr., Phil (2005a), 'Rehabilitating Introspection: A Procedure for a First Person Psychical Science', first presented before the *Association for Politics and the Life Sciences*, Washington, D. C. Available at www.rationology.net
18. Roberts, Jr., Phil (2005b), 'Feelings of Worthlessness: An Annotated Outline of a Theory of Emotional Instability', first presented before the *Association for Politics and the Life Sciences*, Washington, D. C. Available at www.rationology.net
19. Roberts, Jr., Phil (2006), 'Why We Turned Out Like Captain Kirk', first presented before the *Association for Politics and the Life Sciences*. Available at www.rationology.net
20. Sidgwick, Henry (1907), *The Methods of Ethics*.
21. Slovic, Paul; Fischhoff, Baruck and Lichtenstein, Sarah (1976), 'Cognitive Processes and Societal Risk Taking', in Carroll and Payne, 1976.
22. Stein, Edward (1996), *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*.
23. Stich, Stephen P. (1985), 'Could Man Be An Irrational Animal?', *Synthese*, 64, pp.115-135.
24. Stich, Stephen P. (1981), continuing commentary in Cohen, 1981.
25. Tullberg, Jan and Tullberg, Brigitta (2001), 'A Critique of the Naturalistic Fallacy Thesis', *Politics and The Life Sciences*, 20 (2), 165-174.
26. Tversky, A. and Kahneman, D. (1974), 'Judgment Under Uncertainty: Heuristics and Biases', *Science* 185, 1124-31.